

Gefälschte Interviews: Ergebnisse der Interviewerkontrolle beim ALLBUS 1994

Koch, Achim

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Koch, A. (1995). Gefälschte Interviews: Ergebnisse der Interviewerkontrolle beim ALLBUS 1994. *ZUMA Nachrichten*, 19(36), 89-105. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-208984>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

GEFÄLSCHTE INTERVIEWS: ERGEBNISSE DER INTERVIEWERKONTROLLE BEIM ALLBUS 1994

Achim Koch

In der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) 1994 konnte durch den Einsatz eines neuen Stichprobenverfahrens die korrekte Realisierung der 3.505 Interviews systematisch überprüft werden. In sechs Prozent aller Fälle ergab sich ein Verdacht auf Unregelmäßigkeiten. In der Nachkontrolle erwiesen sich ungefähr die Hälfte dieser Interviews als gefälscht. Die von manchen Kritikern der Umfrageforschung geäußerte Behauptung, daß bei mündlichen Umfragen ein großer Teil der Interviews nicht korrekt durchgeführt würde, ist damit zumindest im Falle des ALLBUS nicht zutreffend. Dennoch sollten in Zukunft die Routinekontrollen der Umfrageinstitute verbessert werden, da die gefälschten Interviews in der zuvor durchgeführten Postkartenkontrolle des Umfrageinstituts nicht entdeckt worden waren.

In the German General Social Survey (ALLBUS) 1994 a new sampling design offered the opportunity to control the correct realisation of the 3.505 interviews systematically. In six percent of the cases irregularities were detected. As a result of the examination of these cases, about half of them turned out to be faked by the interviewers. This percentage is much lower than that some critics think is the rule in survey research. Nevertheless, this result demonstrates the need to improve standard control techniques of the survey institutes, because the faked interviews have not been detected by the regular postcard verification.

1. Einleitung

Für die Qualität von Umfrageergebnissen spielt die Phase der Datenerhebung eine entscheidende Rolle. In der Bundesrepublik wird bei mündlichen Umfragen mit bevölkerungsrepräsentativen Stichproben die Feldarbeit überwiegend von kommerziellen Umfrageinstituten erledigt, da nur diese über einen bundesweiten Stab von Interviewern verfügen. Im Rahmen ihrer Tätigkeit als freie Mitarbeiter ist es die Aufgabe der Inter-

viewer, Auswahlpläne und Interviews in der vorgeschriebenen Weise zu realisieren. Abweichungen der Interviewer von den ihnen vorgegebenen Regeln und ihrer Rolle als Interviewer sind in vielerlei Hinsicht möglich: Sie reichen von Veränderungen der Frageformulierungen, mangelhaftem Notieren der Antworten auf offene Fragen, bewertenden Kommentierungen von Äußerungen der Befragten bis zur teilweisen oder vollständigen Fälschung von Interviews.¹⁾ Insbesondere das publikumswirksame Thema der Interviewfälschung wurde jüngst auch in der Presse aufgegriffen (vgl. Der Spiegel 1994). Da wird von Interviewern berichtet, die die Fragebögen zu Hause am Schreibtisch selbst ausfüllen und aufgrund der unzureichenden Kontrollen durch die Umfrageinstitute nur selten entdeckt werden. Es wird die Behauptung aufgestellt, daß bei mündlichen Umfragen ein großer Teil der Interviews nicht korrekt durchgeführt würde, so daß die Gültigkeit der Umfrageergebnisse zweifelhaft sei. Zur Stützung dieser These werden Bekenntnisse einzelner Interviewer herangezogen, die von ihren „erfolgreichen“ - d. h. unentdeckt gebliebenen - Fälschungen berichten (vgl. Dorroch 1994). An verlässlichen quantitativen Daten über das Ausmaß gefälschter Interviews dagegen fehlt es. Veröffentlichungen von Umfrageinstituten, in denen die Ergebnisse ihrer Interviewerkontrollen systematisch ausgewertet werden, liegen nicht vor - vermutlich aufgrund der Besorgnis der Institute, dies könnte als Eingeständnis von Problemen bei der Datenerhebung interpretiert werden.

In der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) 1994 ergab sich durch den Einsatz einer Gemeindestichprobe mit Ziehung von Personenadressen aus den Einwohnermelderegistern eine gute Möglichkeit zur Kontrolle der Feldarbeit. Durch die explizite Vorgabe der zu befragenden Personen an die Interviewer und die Verfügbarkeit von Zusatzinformationen über die Zielpersonen aus den Einwohnermelderegistern konnte die Frage der Fälschung von Interviews systematisch überprüft werden.

2. Formen der Fälschung und Interviewerkontrollen

Bei mündlichen Umfragen können drei Formen der Fälschung unterschieden werden.²⁾ Am schwerwiegendsten sind die sogenannten Totalfälschungen, bei denen der Interviewer den Fragebogen komplett selbst ausfüllt, ohne einen Kontakt zum Befragten herzustellen. Eine Teilfälschung - oder euphemistischer ein „Kurzinterview“ - liegt vor, wenn der Interviewer einige Basisinformationen tatsächlich erfragt und den Rest der Angaben nach seinen Vorstellungen und Annahmen über das Antwortverhalten der Befragten ergänzt.³⁾ In der dritten Variante der Fälschung - der bewußten Verletzung von Auswahlregeln - führt der Interviewer zwar ein vollständiges Interview durch, statt der eigentlichen Zielperson befragt er jedoch absichtlich eine andere Person, weil diese leichter erreichbar oder kooperationsbereiter ist.

Zur Kontrolle der Feldarbeit bei mündlichen Umfragen versenden die Umfrageinstitute in der Regel an eine Stichprobe der mutmaßlichen Befragten Postkarten mit der Bitte um Rückmeldung. Auf diesen Kontrollkarten sollen die betreffenden Personen mitteilen, ob sie tatsächlich interviewt worden sind, und falls dies bejaht wird, sollen einige Fragen zur Durchführung des Interviews beantwortet werden. Erfasst wird dabei u. a. die Interviewdauer, das Alter und Geschlecht der befragten Person sowie Angaben darüber, ob der Interviewer verschiedene Befragungshilfen wie Listen oder Kartenspiele vorgelegt hat. Der Anteil der Interviews einer Studie, der auf diese Weise überprüft wird, kann bei Auftragsvergabe an das Institut individuell vereinbart werden. Standardmäßig werden von den Instituten zumeist nicht mehr als ein Viertel aller Interviews einbezogen.⁴⁾

Mit dieser Postkartenkontrolle ist es theoretisch möglich, sowohl Total- als auch Teilfälschungen von Interviews zu identifizieren. Ob die Verletzung von Auswahlregeln durch die Befragung einer falschen Person entdeckt werden kann, hängt von dem verwendeten Stichprobenverfahren ab. Bei dem sowohl für Studien der Markt- als auch der Sozialforschung am häufigsten eingesetzten Stichprobensystem des Arbeitskreises Deutscher Marktforschungsinstitute („ADM“-Stichprobensystem, vgl. von der Heyde/Löffler 1993; Arbeitsgemeinschaft ADM-Stichproben und Bureau Wendt 1994) wird ein entsprechendes Fehlverhalten in der Regel nicht auffallen, da bei diesem dreistufigen Auswahlverfahren in der zweiten und dritten Stufe die Interviewer die Zielhaushalte und Zielpersonen - entsprechend vorgegebener Regeln, die eine Zufallsauswahl sicherstellen sollen - selbst bestimmen.⁵⁾ Halten die Interviewer diese Auswahlregeln nicht ein, führen aber das Interview mit der falschen Person korrekt durch, dann kann dies durch die Postkartenkontrolle nicht entdeckt werden. Aber auch die Identifizierung von Total- und Teilfälschungen gelingt in der Praxis nur unvollkommen. Der gravierendste Mangel der Postkartentechnik besteht darin, daß nur von einem Teil der angeschriebenen Befragten eine Rückmeldung an das Umfrageinstitut erfolgt. Ein Experiment in den USA (vgl. Hauck 1969) ergab, daß insbesondere Personen, mit denen kein Interview durchgeführt worden ist, die Kontrollkarte nicht zurücksenden. Darüber hinaus erwiesen sich die Informationen auf den zurückgesandten Karten als nicht besonders zuverlässig. So gab z. B. ein Teil der nicht-befragten Personen dennoch an, befragt worden zu sein. Die Schätzung der Interviewdauer durch die Befragten fiel häufig nicht korrekt aus, so daß auch die Identifizierung der teilweise gefälschten Interviews mit Hilfe dieser Angaben nicht unproblematisch ist.⁶⁾

3. Die Interviewerkontrollen beim ALLBUS 1994

Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften wird seit 1980 regelmäßig alle zwei Jahre durchgeführt.⁷⁾ In der Erhebung des Jahres 1994 wurde statt des bislang verwendeten ADM-Stichprobenverfahrens eine Gemeindestichprobe mit Ziehung von Personenadressen aus den Einwohnermelderegistern eingesetzt (vgl. Koch u. a. 1994: 52ff.). In einer zweistufigen, disproportional geschichteten Zufallsauswahl wurden in der ersten Stufe 104 Gemeinden in Westdeutschland und 47 in Ostdeutschland ausgewählt. In der zweiten Stufe wurden aus den Einwohnermelderegistern der ausgewählten Gemeinden jeweils 40 Personen zufällig gezogen; in vier Großstädten in West- und zwei in Ostdeutschland, die aufgrund ihrer Größe mehrmals in die Stichprobe gelangten, ein entsprechendes Vielfaches davon. Mit der Durchführung der Studie war die Infratest Sozialforschung GmbH in München betraut.

Das neue Stichprobenverfahren weist im Vergleich zum ADM-System verschiedene Vorteile auf, sowohl was den Stichprobenplan als auch was die Stichprobenrealisierung im Rahmen der Feldarbeit betrifft (vgl. Koch u. a. 1994: 54ff.). Eine der wichtigsten Verbesserungen besteht in der Reduzierung des Interviewereinflusses bei der Stichprobenbildung. Anders als beim ADM-Stichprobensystem, bei dem die beiden Phasen der Stichprobenbildung und Interviewrealisierung nicht vollständig getrennt sind, liegt bei dem neuen Verfahren vor Feldbeginn eine wohldefinierte Bruttostichprobe mit den Adressen der zu befragenden Personen vor. Die Interviewer erhalten vom Umfrageinstitut die Namen und Adressen der Personen vorgegeben, mit denen sie ihre Interviews durchführen sollen.

Die explizite Vorgabe der Befragungspersonen an die Interviewer bietet im Zusammenspiel mit den - bei nahezu allen Gemeinden - aus den Einwohnermelderegistern zusätzlich verfügbaren Informationen zum Geschlecht und zum Alter der Zielpersonen eine Möglichkeit zur Kontrolle der Feldarbeit, die bei Studien nach dem ADM-Stichprobensystem nicht gegeben ist: Durch einen Vergleich der Informationen aus den Einwohnermelderegistern mit den entsprechenden Angaben aus den Interviews kann die korrekte Realisierung der Interviews überprüft werden. Da bei Totalfälschungen oder der Befragung einer falschen Person insbesondere die Altersangabe im Interview in der Regel nicht mit der Information aus dem Einwohnermelderegister übereinstimmen wird, können diese Fälle relativ zuverlässig identifiziert werden. Teilfälschungen von Interviews sind dagegen auch mit Hilfe dieser Daten nur eingeschränkt zu entdecken - nämlich dann, wenn die Interviewer es unterlassen, während ihres „Kurzinterviews“ das exakte Alter der Zielperson zu ermitteln.

Die wichtigsten Informationen zur Durchführung der ALLBUS-Erhebung 1994 sind in Tabelle 1 zusammengestellt. An der Realisierung der 3.505 Interviews waren insgesamt 693 Interviewer von Infratest beteiligt. Von diesen konnten 609 mindestens ein Interview erzielen, 84 Interviewer blieben ohne Interviewerfolg. Die Ausschöpfung entsprach mit ungefähr 55 Prozent weitgehend den Ergebnissen der Erhebungen der Jahre 1991 und 1992, die Feldzeit dagegen war mit 15 Wochen in West- und elf Wochen in Ostdeutschland die längste in der bisherigen Geschichte des ALLBUS.

Tabelle 1: Ausgewählte Informationen zur Durchführung des ALLBUS 1994*

	West		Ost		Gesamt	
Feldzeit	3.2.-18.5.94 = 15 Wochen		3.2.-21.4.94 = 11 Wochen			
Ausschöpfungsquote	54,1 %		55,9 %		54,7 %	
Realisierte Interviews	2383		1122		3505	
Realisierte Interviews je Interviewer	N	%	N	%	N	%
0	43	9,7	41	16,5	84	12,1
1-2	95	21,4	38	15,3	133	19,2
3-4	99	22,3	49	19,7	148	21,4
5-9	134	30,2	106	42,6	240	34,6
10 und mehr	73	16,4	15	6,0	88	12,7
Gesamt	444	100,0	249	100,1	693	100,0

* Alle Angaben beziehen sich auf den Stand *nach* Durchführung der Routinekontrolle, aber *vor* Durchführung der zusätzlichen Nachkontrolle (s. u.).

Die Arbeit der Interviewer des ALLBUS 1994 war von Infratest routinemäßig mit einer Postkartenkontrolle überprüft worden. Wie in den vorangegangenen Jahren fanden sich dabei keine Hinweise auf größere Unregelmäßigkeiten bei der Durchführung der Umfrage. Lediglich 15 Interviews waren von Infratest aus dem Datensatz entfernt worden, weil sich starke Zweifel an ihrer korrekten Realisierung ergeben hatten.⁸⁾

Tabelle 2: Ergebnisse des Abgleichs zwischen den Angaben zu Geschlecht und Alter im Interview und den entsprechenden Daten aus den Einwohnermeldeämtern

	Realisierte Interviews	Fehlende Angaben	Abgleichbare Fälle	Interviews mit differierenden Angaben	
	N	N	N	N	%
West	2383				
Geschlecht		4	2379	12	0,5
Alter		50	2333	285	12,2
Diff. 1 Jahr				99	4,2
2 Jahre				30	1,3
3 u. mehr Jahre				156	6,7
Ost	1122				
Geschlecht		0	1122	4	0,4
Alter		105	1017	114	11,2
Diff. 1 Jahr				72	7,1
2 Jahre				7	0,7
3 u. mehr Jahre				35	3,4
Gesamt	3505				
Geschlecht		4	3501	16	0,5
Alter		155	3350	399	11,9
Diff. 1 Jahr				171	5,1
2 Jahre				37	1,1
3 u. mehr Jahre				191	5,7

Daß durch diese Standardkontrolle die Regelverletzungen der Interviewer jedoch nur unvollkommen erfaßt werden, zeigte die bei ZUMA zusätzlich durchgeführte Überprüfung.⁹⁾ Unmittelbar im Anschluß an die Lieferung der Daten durch Infratest wurden die Angaben zu Geschlecht und Alter der Zielpersonen aus den Interviews mit den entsprechenden Informationen aus den Einwohnermelderegistern verglichen (vgl. Tab. 2). Erwartungsgemäß erwies sich beim Geschlecht der Anteil differierender Angaben mit circa 0,5 Prozent als sehr gering, denn auch bei inkorrektem Vorgehen konnten die

Interviewer normalerweise dieses Merkmal aus dem Vornamen der Zielperson erschließen. Beim Alter, bei dem eine entsprechende Möglichkeit nicht gegeben war, stimmten dagegen in 12,2 (West) bzw. 11,2 (Ost) Prozent aller Fälle die Angaben aus den Interviews nicht exakt mit den Daten aus den Einwohnermelderegistern überein. Da nicht genau bekannt war, auf welche Weise das Alter in den einzelnen Gemeinden jeweils berechnet worden ist, erschien es jedoch nicht angebracht, geringfügige Unterschiede als Indiz für einen Fehler zu interpretieren.¹⁰⁾ Toleriert man Abweichungen von einem bzw. zwei Jahren, dann reduziert sich der Anteil der auffälligen Interviews deutlich. Mit 6,7 Prozent lag der verbleibende Anteil der Interviews, die eine klärungsbedürftige Differenz von drei oder mehr Jahren beim Merkmal „Alter“ aufwiesen, im Westen ungefähr doppelt so hoch wie im Osten mit 3,4 Prozent.

4. Die Häufigkeit von Interviewerfehlern

Berücksichtigt man die Informationen zu beiden Merkmalen gemeinsam, dann lagen insgesamt 196 Fälle vor, bei denen entweder das Geschlecht nicht übereinstimmte und/oder eine Differenz beim Alter von drei oder mehr Jahren zu konstatieren war (vgl. Tab. 3). Bezogen auf sämtliche 3.505 Interviews entspricht dies einem Anteil von 5,6 Prozent (6,7 Prozent im Westen und 3,2 Prozent im Osten). Auch wenn - wie in Abschnitt 5 noch erläutert wird - nicht in jedem Einzelfall davon ausgegangen werden kann, daß für die Abweichung eine Fälschung oder zumindest ein Fehler des Interviewers verantwortlich ist, soll im folgenden zunächst analysiert werden, welche Interviewer zu den festgestellten Inkonsistenzen beigetragen haben.

Die 196 Fälle mit differierenden Angaben stammen von insgesamt 105 Interviewern (vgl. Tab. 3). Dies bedeutet, daß 17,2 Prozent aller Interviewer (mit mindestens einem realisierten Interview) mindestens ein auffälliges Interview aufweisen. Bei der großen Mehrheit der Interviewer (82,8 Prozent) ist dagegen keinerlei Unregelmäßigkeit zu verzeichnen. In Westdeutschland ist der Anteil der von Abweichungen betroffenen Interviewer mit 21,2 Prozent ungefähr doppelt so hoch wie in Ostdeutschland mit 9,6 Prozent.

Bei 71 der 105 Interviewer (11,7 Prozent aller Interviewer) ist lediglich eines der von ihnen durchgeführten Interviews von einer solchen Inkonsistenz betroffen. Bei 28 weiteren Interviewern (4,6 Prozent aller Interviewer) dagegen weisen zwischen zwei und vier Interviews und bei sechs Interviewern (Ein Prozent aller Interviewer) zwischen fünf und dreizehn Interviews Unregelmäßigkeiten auf. Da im Durchschnitt von den 105 Interviewern mit mindestens einem auffälligem Interview knapp neun (West) bzw. sieben (Ost) Interviews realisiert wurden, erweist sich im Mittel ungefähr jedes fünfte (West) bzw. vierte (Ost) Interview dieser Interviewer als problematisch.

Tabelle 3: Ergebnisse auf Ebene der Interviews und der Interviewer

	West	Ost	Gesamt
Interviews			
1) Realisierte Interviews	2383	1122	3505
2) Interviews mit differierenden Angaben	160	36	196
3) Anteil d. Interviews mit diff. Angaben (2/1)	6,7 %	3,2 %	5,6 %
Alle Interviewer			
4) Interviewer mit mindestens 1 realisierten Interview	401	208	609
5) Interviewer mit mind. 1 Interview mit diff. Angaben	85	20	105
6) Anteil der Interviewer mit mind. 1 Interview mit diff. Angaben (5/4) dabei: Zahl der Interviews mit diff. Angaben	21,2 %	9,6 %	17,2 %
1	14,2 %	6,7 %	11,7 %
2-4	5,7 %	2,4 %	4,6 %
5-13	1,2 %	0,5 %	1,0 %
Interviewer mit mindestens 1 auffälligen Interview			
7) Durchschnittl. Zahl auffälliger Interviews (2/5)	1,9	1,8	1,9
8) Durchschnittl. Zahl durchgeführter Interviews	8,7	6,8	8,3
9) Anteil der auffälligen Interviews (7/8) ¹¹⁾	21,7 %	26,7 %	22,5 %

Tabelle 4: Auftreten von Inkonsistenzen nach demographischen Merkmalen der Interviewer

	West		Ost	
Interviewermerkmal	Realisierte Interviews	Interviews mit diff. Angaben	Realisierte Interviews	Interviews mit diff. Angaben
	N	%	N	%
Geschlecht	n.s.		p<.001	
Männer	1495	7,2	860	2,1
Frauen	887	6,0	260	6,9
Alter	p<.001		p<.001	
18-29 Jahre	154	16,9	47	12,8
30-39 Jahre	371	11,1	93	2,2
40-49 Jahre	599	5,7	187	2,7
50-59 Jahre	681	5,4	399	4,5
60 Jahre und älter	577	3,8	394	1,3
Bildung	p<.01		p<.01	
bis Hauptschulabschl.	587	4,9	96	4,2
Realschulabschl.	960	6,7	262	1,9
Fachh./Abitur	418	10,3	97	9,3
Fachh./Hochschulab.	417	5,8	665	2,7
Gesamt	2383	6,7	1122	3,2

Wie eine Studie der amerikanischen Zensusbehörde ergeben hat, besteht zwischen der Beschäftigungsdauer als Interviewer und der Häufigkeit und Form des Fehlverhaltens ein Zusammenhang: je länger die Beschäftigungsdauer, desto geringer der Anteil gefälschter Interviews, aber desto „elaborierter“ auch die verwendeten Fälschungsmethoden (vgl. Schreiner u. a. 1988). Aufgrund des Fehlens entsprechender Informationen kann im vorliegenden Fall diese Beziehung nicht direkt überprüft werden. Als Beitrag zur Analyse der möglichen Determinanten des Fälschungsverhaltens kann lediglich untersucht werden, ob das Auftreten von Unregelmäßigkeiten mit verschiedenen demographischen Merkmalen der Interviewer in Verbindung steht (vgl. Tab. 4). In Ostdeutschland sind die Interviews von Frauen häufiger als die von Männern von Inkonsistenzen betroffen, in Westdeutschland dagegen läßt sich kein Unterschied zwischen den Geschlechtern ausmachen. Bei der Aufgliederung nach Alterskategorien ergeben sich in beiden Teilen

Deutschlands von allen drei Merkmalen die größten Unterschiede: Die Interviewer unter 40 (West) bzw. unter 30 (Ost) Jahren weisen überproportional viele Interviews mit Abweichungen auf. Es ist nicht ausgeschlossen, daß dieses Ergebnis teilweise den oben berichteten Zusammenhang zwischen der Beschäftigungsdauer und dem Fälschungsverhalten widerspiegelt, da die jüngeren Interviewer im Durchschnitt vermutlich kürzere Zeit als Interviewer tätig sind als ihre älteren Kollegen. Was die Bildung betrifft, ist das auffälligste Ergebnis der hohe Anteil betroffener Interviews bei den Interviewern mit Abitur. Berücksichtigt man zusätzlich das Alter, dann zeigt sich, daß insbesondere die jüngeren Interviewer mit Abitur überproportional viele auffällige Interviews aufweisen. Diese festgestellten Unterschiede in der Fehlerhäufigkeit legen es nahe, bei Kontrollen bestimmten Kategorien von Interviewern besondere Aufmerksamkeit zu widmen.

5. Typen von Interviewfehlern

Um die Gründe aufzuklären, die zu den festgestellten Inkonsistenzen zwischen den Angaben im Interview und den Daten aus den Einwohnermelderegistern geführt haben, wurden von Infratest alle 196 Fälle in eine zusätzliche Nachkontrolle einbezogen, in der versucht wurde, alle betroffenen Zielpersonen telefonisch oder schriftlich sowie in Ausnahmefällen auch persönlich (durch die Chefinterviewer) zu kontaktieren. Zusätzlich wurden die beteiligten Interviewer um eine Stellungnahme gebeten.

In 50 Fällen blieb die Nachprüfung ohne Ergebnis, weil die Zielperson nicht erreicht werden konnte. Bei den verbleibenden 146 Fällen ließen sich fünf verschiedene Fehlertypen als Ursachen der Inkonsistenzen unterscheiden. Insgesamt 45 Interviews (33 im Westen und 12 im Osten) stellen mit großer Sicherheit Totalfälschungen dar, da laut Zielperson bei diesen „Interviews“ keinerlei Kontakt mit dem Interviewer stattgefunden hat. Die Mehrzahl der betroffenen Interviewer bestätigte auf Nachfragen die vorgenommene Fälschung.

51mal wurde zwar ein Interview durchgeführt, jedoch nicht mit der vorgesehenen Zielperson. In der überwiegenden Mehrzahl dieser Interviews wurde statt der Zielperson eine andere Person des Zielhaushalts befragt. Die Auswahl der „Ersatzbefragten“ weist dabei ein bestimmtes Muster auf: Häufig wurde statt der Tochter oder des Sohnes die im gleichen Haushalt lebende Mutter oder der Vater interviewt. Dieses Vorgehen hat zur Folge, daß das Durchschnittsalter bei den eigentlich zu befragenden 51 Personen nach den Daten der Einwohnermelderegister deutlich geringer ist als das der tatsächlichen Befragten nach den Angaben im Interview: 39 versus 50 Jahre. Hinzuweisen ist im übrigen darauf, daß bei circa einem Drittel der 51 Interviews das Fehlverhalten dem Interviewer kaum anzulasten ist. Laut dem Ergebnis der Nachkontrolle lebten in den betreffenden Haushalten

mehrere Personen, bei denen nicht nur der Nachname, sondern auch der Vorname oder Teile des Vornamens übereinstimmten, so daß in diesen Fällen die Interviewer die Entscheidung über die zu befragende Person selbst zu treffen hatten.

Tabelle 5: Ergebnis der Überprüfung der 196 Fälle mit differierenden Angaben

Fehlertyp	Betroffene Interviews		Beteiligte Interviewer			
			Zahl der Interviewer	Durchschnittl. Zahl betr. Interviews (1/3)	Durchschnittl. Zahl realisierter Interviews	Anteil fehlerhafter Interviews (4/5)
	N (1)	% (2)	N (3)	N (4)	N (5)	% (6)
1. Totalfälschung	45	30,8	6	7,5	13,0	57,7
2. Falsche Person befragt	51	34,9	42	1,2	8,2	14,8
3. Zielperson bestätigt Interview und Angaben des Einwohnermeldeamtes	31	21,2	25	1,2	8,2	15,1
4. Zielperson bestätigt Interview und Angaben im Interview	9	6,2				
5. Technische Fehler	10	6,8				
Gesamt Überprüfung ohne Ergebnis	146 50	99,9				

In 31 Fällen wurde die Realisierung des Interviews und die Richtigkeit der Angaben des Einwohnermeldeamtes durch die Zielperson bestätigt. Als Ursache für die aufgetretenen Abweichungen kommen hier Fehler bei der Interviewdurchführung sowohl auf seiten des Befragten als auch auf seiten des Interviewers bzw. Mängel in der Datenaufnahme und nachfolgenden Datenaufbereitung in Betracht. Eine genauere Inspektion ergab, daß eine Ersetzung der bisherigen Angaben im Interview durch die von den Zielpersonen bestätigten Alters- und Geschlechtsangaben der Einwohnermeldeämter nur bei einem Teil der 31 Fälle sinnvoll und ohne Probleme möglich war. In der Mehrzahl der Fälle dagegen hätten die korrigierten Werte zu Widersprüchen mit den sonstigen Daten im Interview - insbe-

sondere den Informationen zum Zeitpunkt der ersten Erwerbstätigkeit sowie den Angaben zur Aufenthaltsdauer im jeweiligen Bundesland - geführt. Aus diesem Grund kann keine eindeutige Bewertung dieses Fehlertyps erfolgen. Nicht ausgeschlossen ist z. B., daß bei einem Teil dieser Interviews die geschilderten Probleme daraus resultieren, daß die Interviewer lediglich ein „Kurzinterview“ durchgeführt und den Rest des Fragebogens (inclusive der Altersangabe) selbst ausgefüllt haben.¹²⁾

Mit Sicherheit kein Interviewerfehler liegt dagegen bei den verbleibenden 19 Fällen vor. Von neun Zielpersonen wurden das Interview und die entsprechenden Angaben im Fragebogen bestätigt. Als Ursache für die Diskrepanzen kommen hier Fehler in den von den Einwohnermeldeämtern übermittelten Daten bzw. Mängel bei deren Aufbereitung in Betracht. Bei zehn Interviews waren die Inkonsistenzen die Folge technischer Fehler bei der Organisation der Umfrage. Aufgrund der Vertauschung von Identifikationsnummern wurden bei diesen Interviews für den Datenabgleich fälschlicherweise die Angaben unterschiedlicher Personen einander zugeordnet.

Insgesamt ergab die Nachprüfung bei den 146 kontaktierten Zielpersonen damit, daß nur in circa der Hälfte der inkonsistenten Fälle definitiv eine Fälschung durch die Interviewer erfolgte - sei es, daß sie die betreffenden Fragebögen komplett selbst ausgefüllt oder das Interview absichtlich mit einer anderen Person durchgeführt haben. In circa 35 Prozent aller Fälle, bei denen aufgrund von Namensgleichheit eine falsche Person befragt bzw. das Interview und die Angaben des Einwohnermeldeamtes bestätigt wurden, ist der Fehler dem Interviewer nicht anzurechnen bzw. nicht direkt belegbar. Eindeutig keinen Interviewerfehler stellen die 13 Prozent der Interviews dar, bei denen die Abweichungen die Folge von Fehlern in den Einwohnermeldedaten bzw. technischer Probleme sind.

Für die Fehlertypen, bei denen ein Verschulden des Interviewers sicher vorliegt bzw. nicht ausgeschlossen werden kann, sind in Tabelle 5 noch einige Kennziffern auf Interviewerebene wiedergegeben. Anhand dieser kann man deutliche Unterschiede zwischen den Totalfälschungen auf der einen und den beiden anderen Fehlertypen auf der anderen Seite erkennen. Interviewer, die Totalfälschungen vornehmen, fälschen in der Regel nicht nur ein Interview, sondern mehrere. Beim ALLBUS 1994 haben diese Interviewer im Durchschnitt 7,5 Interviews gefälscht. Die Zahl der von ihnen insgesamt durchgeführten Interviews liegt mit durchschnittlich 13 dementsprechend hoch. Zum Vergleich: Der Mittelwert aller 609 Interviewer (mit mindestens einem realisierten Interview) liegt bei 5,8, der Mittelwert der 105 Interviewer mit mindestens einem auffälligen Interview bei 8,3 Interviews. Allerdings haben auch die „Fälscher“ nicht sämtliche ihrer Interviews unlauter fabriziert. Im Durchschnitt haben sie etwas mehr als die Hälfte ihrer Interviews gefälscht, lediglich einer der sechs Interviewer hat sämtliche Interviews komplett selbst

ausgefüllt. Die Regelverletzung durch die Befragung einer falschen Person sowie der nicht eindeutig zu klärende Fehlertyp dagegen sind - was die Auftretenshäufigkeit betrifft - stärker punktueller Natur. Im Durchschnitt ist bei den beteiligten Interviewern lediglich ein Interview von diesem Fehler betroffen. Der durchschnittliche Anteil der auffälligen Interviews an allen jeweils realisierten Interviews beträgt denn hier auch nur circa 15 Prozent.

6. Zusammenfassung und Schlußfolgerung

Die systematische Kontrolle der Datenerhebung beim ALLBUS 1994 ergab bei knapp sechs Prozent der 3.505 Interviews einen Verdacht auf Unregelmäßigkeiten. In der einzelfallbezogenen Nachprüfung erwiesen sich ungefähr die Hälfte dieser Fälle - d. h. drei Prozent aller Interviews - als gefälscht. Auch wenn der tatsächliche Anteil der Fälschungen vermutlich noch etwas höher anzusetzen ist, weil durch die Kontrolle Fälschungen in Form sog. „Kurzinterviews“ nur in beschränktem Maße identifiziert werden konnten, belegt dieses Ergebnis, daß zumindest beim ALLBUS die in der Presse kolportierte Aussage „Jedes dritte Interview ist gefälscht“ (vgl. Der Spiegel 1994) keinesfalls zutrifft. In den von Schnell (1991) durchgeführten Untersuchungen wirkten sich Fälschungen, die fünf Prozent eines Datensatzes ausmachten, weder auf die Berechnung univariater Statistiken noch auf multivariate Analysen aus. Die Gefahr, daß durch Fälschungen von Interviewern die Ergebnisse von Analysen mit ALLBUS-Daten beeinträchtigt werden, ist deshalb kaum gegeben.¹³⁾

Trotz der relativ gesehen geringen Zahl gefälschter Interviews legt die durchgeführte Untersuchung nahe, in Zukunft über verbesserte Formen der Interviewerkontrolle nachzudenken. Die Überprüfung durch ZUMA fand im Anschluß an die Routinekontrolle des Umfrageinstitutes statt, bei der 15 gefälschte Interviews entdeckt und aus dem Datensatz entfernt worden waren. Dennoch konnten in der Nachkontrolle weitere 45 Totalfälschungen sowie 51 Fälle, bei denen die falsche Person befragt wurde, identifiziert werden. Dies zeigt, daß die als Standardtechnik eingesetzte Postkartenkontrolle nur ein sehr unvollständiges Bild von dem tatsächlichen Fehlverhalten der Interviewer liefert.

Anmerkungen

- 1) Für das Auftreten dieser Fehler spielen auf seiten der Interviewer Uninformiertheit, Nachlässigkeit oder das Streben nach „Erleichterung“ ihrer Arbeit und Aufbesserung ihres Honorars eine Rolle. Es wäre jedoch sicherlich nicht angemessen, sämtliche Fehler alleine dem wie auch immer gearteten Versäumnis der Interviewer anzulasten. Auch die Umfrageinstitute und die Auftraggeber der Studien tragen durch mangelhafte Schulungen, unverständliche und zu lange Fragebögen sowie unzureichende Bezahlung zu Interviewerfehlern bei.
- 2) Die drei genannten Formen der Fälschung - Total- und Teilfälschung sowie Verletzung von Auswahlregeln - stellen die wichtigsten Fälschungsvarianten bei mündlichen Umfragen dar. Weitere - im folgenden nicht berücksichtigte - Fälschungsmöglichkeiten betreffen die bewußte Fehlklassifikation von nichterreichten oder nichtkooperativen Zielpersonen als stichprobenneutrale Ausfälle oder das Abweichen vom vorgeschriebenen Durchführungsmodus des Interviews (z. B. telefonisch statt persönlich). Zu den verschiedenen Formen der Fälschung vgl. Schreiner u. a. 1988.
- 3) Die von Reuband (1990) und Schnell (1991) durchgeführten Experimente belegen, daß Teilfälschungen von Interviews auf der Basis einiger bekannter demographischer Merkmale prinzipiell „erfolgreich“ durchgeführt werden können. Die Differenzen zwischen echten und gefälschten Interviews erwiesen sich - sowohl hinsichtlich der Randverteilungen und univariaten Statistiken als auch der Ergebnisse multivariater Analysen - als erstaunlich gering.
- 4) Neben der Zahl der zu kontrollierenden Interviews können für die jeweilige Studie darüber hinaus verschiedene weitere Kriterien mit dem Umfrageinstitut abgesprochen werden, die bei der Auswahl der zu überprüfenden Interviews zu berücksichtigen sind (z. B. mindestens ein Klumpen von jedem Fulltime-Interviewer; alle Interviews von Interviewern, die erstmals eingesetzt werden).
- 5) Bei dem dreistufigen ADM-Stichprobenverfahren wird zunächst eine Stichprobe von Wahlbezirken gezogen. In der zweiten Auswahlstufe ermitteln die Interviewer in den ausgewählten Wahlbezirken - ausgehend von einer zufällig ausgewählten Startadresse und entsprechend bestimmter Begehungsregeln - eine vorgegebene Zahl von Haushalten. In der dritten Ziehungsstufe wird von den Interviewern in den ausgewählten Haushalten mit Hilfe eines vorbereiteten Zufallsschemas („Schwedenschlüssel“) die letztendlich zu befragende Person bestimmt.

- 6) Da Befragte eher dazu neigen, die tatsächliche Interviewdauer zu unter- als zu überschätzen, besteht hier auch die Gefahr, daß Interviewer zu Unrecht einer Teilfälschung verdächtigt werden.
- 7) Mit dem ALLBUS wird das wissenschaftliche Ziel der Untersuchung sozialen Wandels und das praktische Ziel der Verbesserung der Infrastruktur in den Sozialwissenschaften durch die schnelle Bereitstellung von Primärdaten verfolgt (vgl. Braun/Mohler 1991). In jeder Erhebung wird eine Zufallsstichprobe der Bevölkerung mit einem teils konstanten, teils variablen Fragenprogramm befragt. Die Grundgesamtheit der Umfrage bildet seit 1991 die erwachsene Wohnbevölkerung in West- und Ostdeutschland. In sämtlichen Erhebungen zwischen 1980 und 1992 kam das ADM-Stichprobensystem zum Einsatz. Die Stichprobengröße betrug bis 1991 insgesamt 3.000 Befragte, seit 1992 liegt sie bei 3.500 Personen (2.400 West, 1.100 Ost). Die Umfragen der Jahre 1980 bis 1986 sowie die Zusatzstudie 1991 wurden durch Mittel der Deutschen Forschungsgemeinschaft (DFG) gefördert, die verbleibenden Umfragen seit 1988 wurden von Bund und Ländern gemeinsam im Rahmen der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS) finanziert.
- 8) Kontrollkarten wurden an circa 25 Prozent aller Befragten versandt. In Westdeutschland gelangten 596 Interviews von 152 Interviewern in die Kontrolle. Von 429 Befragten (=72 Prozent) wurde die Karte zurückgesandt. In neun Fällen ergaben sich starke Zweifel an der korrekten Realisierung des Interviews, die auch von den kontaktierten Interviewern nicht ausgeräumt werden konnten. In Ostdeutschland wurden in gleicher Weise 287 Interviews von 70 Interviewern überprüft. Dabei erzielte man einen Rücklauf von 173 Antworten (= 60 Prozent). Sechs Fälle erwiesen sich als nicht korrekt durchgeführt.
- 9) Da uns für die von Infratest durchgeführte Postkartenkontrolle lediglich die in Anmerkung 8 geschilderten Informationen vorliegen, ist es bedauerlicherweise nicht möglich, die Ergebnisse beider Kontrollen auf Individualebene miteinander zu vergleichen.
- 10) Die Ziehung der Personenadressen erfolgte in der Mehrzahl der ausgewählten Gemeinden zwischen Oktober und Dezember 1993, bei einem Teil jedoch erst im Januar bzw. Februar 1994. Je nachdem, zu welchem Zeitpunkt die Adressziehung durchgeführt und nach welchem Verfahren das Alter berechnet wurde (wurde nur mit den Jahresangaben gearbeitet oder wurde das exakte Geburtsdatum und der Stichtag der Adressziehung verwendet?), konnte das auf Basis der Einwohnermeldedaten ermittelte Alter um bis zu zwei Jahren variieren. In die zusätzliche Nachkontrolle (vgl. Abschnitt 5) wurden deshalb

nur Fälle mit einer Differenz von drei und mehr Jahren einbezogen, da bei diesen mit Sicherheit irgendeine Unregelmäßigkeit zugrunde liegen mußte. Zwar befinden sich vermutlich auch unter den Fällen mit einer geringeren oder gar keiner Differenz beim Alter einige weitere gefälschte Interviews, doch dürfte deren Anteil zu gering sein, um die geschilderten Ergebnisse substantiell zu verändern.

11) Aufgrund der Unterschiede, die zwischen den Interviewern in der Anzahl der durchgeführten Interviews bestehen, ergeben sich leicht differierende Werte, wenn der Anteil auffälliger Interviews zunächst je Interviewer berechnet und anschließend über alle Interviewer gemittelt wird. Der durchschnittliche prozentuale Anteil auffälliger Interviews und die zugehörige Standardabweichung beträgt bei dieser Berechnungsweise im Westen 25,8 % ($s=20,1$), im Osten 29,7 % ($s=21,0$) und insgesamt 26,6 % ($s=20,3$).

12) Möglich ist jedoch auch, daß in Einzelfällen die Ergebnisse der Nachkontrolle selbst fehlerbehaftet sind.

13) Dies gilt beim ALLBUS 1994 umso mehr, als die 45 Totalfälschungen natürlich aus dem Datensatz entfernt worden sind. Die 51 Fälle, bei denen die Auswahlanweisung des Stichprobenplanes nicht eingehalten und eine falsche Person befragt worden ist, wurden dagegen im Datensatz belassen. Eine Streichung dieser Fälle hätte bedeutet, beim ALLBUS 1994 weit strengere Maßstäbe anzulegen als bei den vorangegangenen Erhebungen nach dem ADM-Stichprobensystem, da bei diesem aufgrund des eingesetzten Random-Route-Verfahrens zur Haushaltsbestimmung und des Schwedenschlüssels zur Auswahl der Zielpersonen ein entsprechendes Fehlverhalten der Interviewer vermutlich in stärkerem - aber eben nicht identifizierbarem - Ausmaß aufgetreten sein dürfte.

Literatur

Arbeitsgemeinschaft ADM-Stichproben und Bureau Wendt, 1994: Das ADM-Stichproben-System. Stand: 1993. S. 188-202 in: S. Gabler/J. H. P. Hoffmeyer-Zlotnik/D. Krebs (Hrsg.), Gewichtung in der Umfragepraxis. Opladen: Westdeutscher Verlag.

Braun, M./Mohler, P. Ph., 1991: Die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS): Rückblick und Ausblick in die neunziger Jahre. ZUMA-Nachrichten 29: 7-28.

Dorroch, H., 1994: Meinungsmacher-Report. Wie Umfrageergebnisse entstehen. Göttingen: Steidl.

Hauck, M., 1969: Is Survey Postcard Verification Effective? Public Opinion Quarterly 23: 117-120.

Von der Heyde, Ch./Löffler, U., 1993: Die ADM-Stichprobe. Planung und Analyse, Heft 5: 49-53.

Koch, A./Gabler, S./Braun, M., 1994: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1994. ZUMA-Arbeitsbericht 94/11. Mannheim.

Reuband, K.-H., 1990: Interviews, die keine sind. „Erfolge“ und „Mißerfolge“ beim Fälschen von Interviews. Kölner Zeitschrift für Soziologie und Sozialpsychologie 42: 706-733.

Schnell, R., 1991: Der Einfluß gefälschter Interviews auf Survey-Ergebnisse. Zeitschrift für Soziologie 20: 25-35.

Schreiner, I./Pennie, K./Newbrough, J., 1988: Interviewer Falsification In Census Bureau Surveys. American Statistical Association. Proceedings of the Section on Survey Research Methods: 491-496.

Der Spiegel, 1994: Ohrfeige an der Haustür. Heft 26: 41-46.